PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/56, 15/61, 15/82, 15/83, 9/40, 9/90, C12P 19/20, 19/24, C08B 37/00, A01H 5/00

A2

(11) International Publication Number: WO 97/20937

(43) International Publication Date: 12 June 1997 (12.06.97)

(21) International Application Number: PCT/EP96/05581

(22) International Filing Date: 2 December 1996 (02.12.96)

(30) Priority Data:
9524752.4 4 December 1995 (04.12.95) GB

(71) Applicant (*for all designated States except US*): DANISCO A/S [DK/DK]; Langebrogade 1, P.O. Box 17, DK-1001 Copenhagen K (DK).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): JØRSBOE, Morten [DK/DK]; Ejegodvej 43, st. tv., DK-4800 Nykøbing F (DK). BRUNSTEDT, Janne [DK/DK]; Dueholm 18, DK-4000 Roskilde (DK). PETERSEN, Steen, Guldager [DK/DK]; Grøndals Parkvej 2B, 3.tv., DK-2720 Copenhagen Vanløse (DK).

(74) Agents: HARDING, Charles, Thomas et al.; D. Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MODIFICATION PROCESS

(57) Abstract

An *in vivo* modification process is described. The *in vivo* modification process affects the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof.

WEST

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

MODIFICATION PROCESS

The present invention relates to a modification process.

In particular, the present invention relates to an *in vivo* modification process.

Galactomannans are a heterogenous group of cell wall polysaccharides consisting of a β-1-4 linked mannan backbone with varying numbers of α-1-6 linked galactose side chains.

The galactomannans of most significant industrial use are obtained from the endosperms of the legumes guar (*Cyamopsis tetragonolobus*) and locust bean (*Ceratonia siliqua*). These galactomannans differ in their galactose content, guar having a galactose to mannose ratio of approximately 1:1.6, whereas the ratio for locust bean gum (LBG) is approximately 1:3.4.

The differences in galactose content have significant effects on the functional properties of guar gum and LBG. Both galactomannans form highly viscous solutions at low concentrations (1-2%) but LBG has the additional property of being able to form firm gels with other polysaccharides such as xanthan, carrageenan and agarose. LBG is extensively used by the food industry in dairy products (notably ice cream), salad dressings, sauces, low calorie products and pet foods. However, the use of LBG is restricted by the high price and irregular supply.

Therefore, large scale production of galactomannans with improved functional properties, such as a result of an increased mannose to galactose ratio (such as that similar to LBG), is desirable.

Due to the generic chemical similarity between guar gum and LBG and the much lower price on guar gum, it has been attempted *in vitro* to convert guar gum into a galactomannan with LBG-like properties and with a chemical composition similar to LBG.

An example of such an *in vitro* treatment includes the use of α-galactosidase. In this regard, see McCleary *et al* 1983 and EP-A-0255153.

By using α-galactosidase purified from guar seeds, guar gum with galactose contents of 10-34% was obtained (Bulpin *et al* 1990). Analysis of the gelation behaviour of the modified guar gum showed that a preparation with a galactose content of 24% formed mixed gels with carrageenan displaying similar rheological properties as LBG. In comparison, the galactose content of untreated guar gum was 38% and 23% for LBG.

However, from an industrial point of view *in vitro* degalactosylation of guar gum is associated with a number of problems.

First, huge amounts of α-galactosidase have to be prepared as about 40% of the galactose in the guar gum has to be removed.

Second, during incubation it is very important that no hydrolysis of the mannan backbone occurs necessitating the use of highly purified α-galactosidase preparations devoid of any trace of mannanase activity. A procedure for heterologous production of the α-galactosidase from guar seed has been published (Overbeeke *et al* 1986). However, the produced α-galactosidases from the tested species were purified before investigating the action on guar gum suggesting that the mannanase problem remains to be solved.

Third, the yield of galactomannan is reduced because a 40% reduction in galactose content corresponds to approximately 15% less modified guar gum. The released galactose may be undesirable in the final product and may have to be removed.

Fourth, there is a considerable risk for depolymerisation of the galactomannan during incubation with α-galactosidase.

Also, there is a risk for contaminating microorganisms to colonise the reaction mixture releasing endo-β-mannanases.

Finally, water has to be removed from the reaction mixture. In addition to the cost of this process, it will also result in concentration of the buffer that may be used for obtaining optimal reaction conditions.

These examples demonstrate that the present methods for the modification of guar gum are associated with problems, some of which are associated with considerable costs.

There is therefore need to have an improved method for the modification of guar gum.

In this regard, we now realise that it would be beneficial if the modification of a mannose/galactose containing compound (such as guar gum) were to occur *in vivo* in plants, such as guar plants, by use of recombinant DNA techniques.

Thus, in its broadest sense, the present invention relates to *in vivo* modification of a mannose/galactose containing compound - such as guar gum - in an organism (or part thereof) capable of synthesising that compound by a method that is not native to that organism - such as by a method that makes use of recombinant DNA techniques. The modification may occur to in relation to any one or more of the precursors of the compound (e.g. mannose and/or galactose) or in relation to the compound itself (i.e. modification of the mannose and/or galactose units of a compound comprising same).

In particular, the present invention relates to an *in vivo* modification process that affects, preferably increases, the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof. This *in vivo* modification process is not a naturally occurring process.

Thus, with the *in vivo* process of the present invention, it is possible to alter the internal *in vivo* ratio of mannose to galactose within an organism and/or the ratio of mannose to galactose of a mannose/galactose compound thereof.

One of the requirements for the production of *in vivo* modified guar gum is the availability of a method for the introduction of suitable genes into guar. This has been accomplished to a limited extent by Jørsboe and Okkels (1994) who transferred a selectable and screenable gene used for the development of the transformation method. These authors did not report on transformation with a gene to affect the mannose to galactose ratio. This is an important point as, from a biotechnological point of view, the major obstacle for the production of *in vivo* modified guar gum is the lack of knowledge of galactomannan biosynthesis. Up until now, no genes or gene products which control the biosynthesis of guar gum *in vivo* have been isolated and characterised. However, we have now determined some of the genes or gene products which control the biosynthesis of guar gum *in vivo* - thus enabling us to modify guar gum *in vivo*.

In one preferred aspect, the present invention relates to an *in vivo* modification process that affects, preferably increases, the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, the *in vivo* modification process comprising expressing a nucleotide sequence coding for a gene product, which gene product has an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the nucleotide sequence is not a natural nucleotide sequence to the organism (or part thereof).

In another preferred aspect, the present invention relates to an *in vivo* modification process that affects, preferably increases, the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, the *in vivo* modification process comprising allowing a gene product that is capable of having an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

to have an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the gene product has not been expressed by a nucleotide sequence that is a natural nucleotide sequence to the organism (or part thereof).

Another broad aspect of the present invention relates to the use of a nucleotide sequence to affect *in vivo*, preferably to increase, the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, wherein the nucleotide sequence encodes a gene product which has an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the nucleotide sequence is not a natural nucleotide sequence to the organism (or part thereof).

Another broad aspect of the present invention relates to the use of a gene product to affect *in vivo*, preferably to increase, the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, wherein the gene product has an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the gene product is not expressed by a nucleotide sequence that is a natural nucleotide sequence to the organism (or part thereof).

The term "mannose/galactose containing compound" means a compound comprising at least one mannose group and at least one galactose group.

In each of these preferred aspects, it is preferred that the mannose/galactose containing compound is galactomannan.

In each of these preferred aspects, it is more preferred that the mannose/galactose containing compound is guar gum.

In each of these preferred aspects, it is more preferred that the organism capable of producing a mannose/galactose containing compound is a guar plant and the mannose/galactose containing compound thereof is galactomannan. However, other galactomannan producing plants are encompassed such as fenugreek and lucerne. Plants that are considered not to produce appropriate quantities of galactomannan belong to the family *Solanacea* and the species *Nicotiana tabacum*.

The term "organism (or part thereof) capable of producing a mannose/galactose containing compound" also includes any suitable organism - in particular a plant - capable of producing a mannose/galactose containing compound, such that the internal *in vivo* ratio of mannose to galactose of that organism is altered. The term also includes any part of an organism that is capable of producing a mannose/galactose containing compound, such that the ratio of mannose to galactose of that part is altered. The term also includes a part when within an organism or in a live culture medium. Preferably, the part is when within an organism *per se*. An example of a part is seed.

The term "a natural nucleotide sequence to the organism" means an entire nucleotide sequence that is in its natural environment and when operatively linked to an entire promoter with which it is naturally associated, which promoter is also in its natural environment.

The term "mannose and galactose precursors" includes mannose *per se* or derivatives thereof and/or galactose *per se* or derivatives thereof as precursors for the biosynthesis of a mannose/galactose containing compound, preferably galactomannan. In addition, the term includes precursors for mannose *per se* or derivatives thereof and/or galactose *per se* or derivatives thereof which in turn are used as precursors for the biosynthesis of a mannose/galactose containing compound, preferably galactomannan. Preferably, the term means mannose *per se* or derivatives thereof (such as mannose-6-phosphate or GDP-mannose) and/or galactose *per se* or derivatives thereof as precursors for the biosynthesis of galactomannan, preferably guar galactomannan.

The term "gene product" includes peptides, polypeptides, proteins, enzymes and RNA. Preferably, the term means an enzyme.

Preferably the *in vivo* mannose-to-galactose ratio of the organism (or part thereof) or mannose/galactose containing compound thereof is higher than that of the guar plant or the galactomannan thereof.

More preferably the *in vivo* mannose-to-galactose ratio of the organism (or part thereof) or mannose/galactose containing compound thereof is substantially similar to that of the locust bean or the galactomannan thereof.

Preferably the organism (or part thereof) or mannose/galactose containing compound thereof is a guar plant or the gum thereof.

The present invention also covers a mannose/galactose containing compound when prepared by the process of the present invention. This mannose/galactose containing compound will be referred to as a mannose/galactose containing compound according to the present invention.

In addition, the present invention also covers a foodstuff comprising a mannose/galactose containing compound according to the present invention.

In addition, the present invention also covers a composition - such as a foodstuff - comprising a mannose/galactose containing compound according to the present invention admixed with another polysaccharide. Preferably that other saccharide is any one or more of xanthan, carrageenan and agarose.

In addition, the present invention covers methods for preparing compositions or foodstuffs according to the present invention comprising mixing the mannose/galactose containing compound according to the present invention with another suitable ingredient.

The broad aspects of the present invention can be achieved by one or more appropriate strategies, wherein each strategy constitutes a preferred embodiment of the present invention.

A first strategy relates to the use of one or more gene products, or nucleotide sequences coding for same, wherein the gene products are useful in the biosynthesis of GDP-mannose. This strategy involves the transformation of one or more of the genes encoding enzymes that are required for the biosynthesis of GDP-mannose - namely the enzyme phosphomannose isomerase (PMI) and/or the enzyme phosphomannose mutase and/or the enzyme GDP-mannose pyrophosphorylase.

In this regard, it is believed that one or more of the gene products that are useful in the biosynthesis of GDP-mannose increase the levels of mannose-6-phosphate, which in turn increase the mannose-to-galactose ratio of a mannose/galactose containing compound, such as a galactomannan.

A preferred aspect of the first strategy relates to at least the use of PMI and/or the nucleotide sequence coding for the same. In this regard, it is believed that the PMI gene product increases the levels of mannose-6-phosphate, which in turn increases the mannose-to-galactose ratio of a mannose/galactose containing compound, such as a galactomannan. It is further preferred that the PMI is a plant PMI.

A second strategy relates to the use of a α-galactosidase enzyme and the nucleotide sequence coding for the same. With this strategy it is possible to utilise α-galactosidase, such as that from senna or from coffee beans, to alter *in vivo* the mannose-to-galactose ratio of a mannose/galactose containing compound, such as a galactomannan.

A third strategy relates to the combination of the first strategy with the second strategy, which strategies can be used in any order or simultaneously.

A preferred aspect of the present invention relates to a construct comprising or expressing any one or more of the nucleotide sequences of the present invention.

Another preferred aspect of the present invention relates to a vector comprising or expressing any one or more of the constructs or nucleotide sequences of the present invention.

Another preferred aspect of the present invention relates to a plasmid comprising or expressing any one or more of the vectors, constructs or nucleotide sequences of the present invention.

Another preferred aspect of the present invention relates to a transgenic organism (or part thereof) comprising or expressing any one or more of the plasmids, vectors, constructs or nucleotide sequences of the present invention.

Other preferred aspects of the present invention include methods of expressing or allowing expression or transforming any one or more of the nucleotide sequences, the constructs, the plasmids, the vectors, the cells, the tissues, the organs or the organisms, as well as the products thereof.

Further preferred aspects of the present invention include uses of the gene products for preparing or treating foodstuffs, including animal feed.

The present invention also relates to isolating a guar gum prepared by the method of the present invention.

The present invention also relates to a guar gum prepared by the method of the present invention.

The first strategy of the present invention will now be described in more detail by way of reference to further preferred aspects of the present invention.

According to a first preferred aspect of this aspect of the present invention there is provided an enzyme comprising the amino acid sequence shown in Figure 1, or a variant, homologue or fragment thereof.

According to a second preferred aspect of this aspect of the present invention there is provided a nucleotide sequence coding for the enzyme of the first aspect or a sequence that is complementary thereto.

According to a third preferred aspect of this aspect of the present invention there is provided a nucleotide sequence comprising the sequence shown in Figure 1, or a variant, homologue or fragment thereof or a sequence that is complementary thereto.

According to a fourth preferred aspect of this aspect of the present invention there is provided a construct comprising or expressing the invention according to any one of the earlier aspects.

According to a fifth preferred aspect of this aspect of the present invention there is provided a vector comprising or expressing the invention of any one of the earlier aspects.

According to a sixth preferred aspect of this aspect of the present invention there is provided a plasmid comprising or expressing the invention of any one of the earlier aspects.

According to a seventh preferred aspect of this aspect of the present invention there is provided a transgenic organism (or part thereof) comprising or expressing the invention according to any one of the earlier aspects.

Preferably in these preferred aspects of this aspect of the present invention the nucleotide sequence or the enzyme is that as defined in or is contained within or is expressed by the above-mentioned aspects of the present invention.

Other preferred aspects of this aspect of the present invention include methods of expressing or allowing expression or transforming any one of the nucleotide sequence, the construct, the plasmid, the vector, the cell, the tissue, the organ or the organism, as well as the products thereof.

Further preferred aspects of this aspect of the present invention include uses of the enzyme for preparing or treating foodstuffs, including animal feed.

A preferred aspect of this aspect of the present invention thus relates to the enzyme phosphomannose isomerase ("PMI") and a nucleotide sequence coding for that enzyme. In particular, the preferred aspect of the present invention relates to recombinant PMI.

In addition, the preferred aspect of the present invention relates to the use of that recombinant PMI to alter the mannose-to-galactose ratio of either an organism (or part thereof) and/or a mannose/galactose containing compound thereof, especially the mannose-to-galactose ratio of galactomannan.

One of the key advantages of the present invention is that by using the recombinant PMI it is possible to increase the mannose-to-galactose ratio of an organism (or part thereof) and/or a mannose/galactose containing compound thereof, in particular *in vivo* modified guar gum. This advantageous aspect is achieved by the insertion of a gene or genes encoding a gene product or products which are involved in the biosynthesis of mannose/galactose containing compounds such as mannose-6-phosphate, which gene is most preferably the nucleotide sequence of the present invention.

Other key advantages are that the recombinant enzyme can be prepared easily and in large quantities. Also, the nucleotide sequence can be used to change the *in vivo* ratio of mannose to galactose levels when inserted (preferably stably inserted) into the genome of an organism (or part thereof).

Preferably the nucleotide sequence is a DNA sequence.

In a highly preferred embodiment the nucleotide sequence is a recombinant DNA sequence.

Preferably the nucleotide sequence is obtainable from deposit NCIMB 40774.

In a highly preferred embodiment the enzyme is expressed by use of recombinant DNA techniques.

Preferably the enzyme is expressed by a nucleotide sequence obtainable from deposit NCIMB 40774.

Preferably the organism is a plant.

More preferably the plant is a guar plant.

Preferably the mannose/galactose containing compound is guar gum.

The enzyme or nucleotide sequence(s) coding for same may be used *in vivo* in combination with one or more other enzymes or nucleotide sequence(s) coding for same, which enzymes or nucleotide sequence(s) coding for same are preferably prepared by use of recombinant DNA techniques. The PMI enzyme or nucleotide sequence(s) coding for same may also be used *in vitro*. The PMI enzyme or nucleotide sequence(s) coding for same may also be used with one or more other enzymes or nucleotide sequence(s) coding for same, which enzymes or nucleotide sequence(s) coding for same are preferably prepared by use of recombinant DNA techniques.

The terms "variant", "homologue" or "fragment" in relation to the enzyme include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) amino acid from or to the sequence providing the resultant amino acid sequence has PMI activity, preferably having at least the same activity of the enzyme shown in Figure 1. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant enzyme has PMI activity. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to the sequence shown in Figure 1. More preferably there is at least 95%, more preferably at least 98%, homology to the sequence shown in the attached Figure 1.

The terms "variant", "homologue" or "fragment" in relation to the nucleotide sequence coding for the enzyme include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence providing the resultant nucleotide sequence codes for an enzyme having PMI activity, preferably having at least the same activity of the enzyme shown in Figure 1. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence codes for an enzyme having PMI activity. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to the sequence shown in Figure 1. More preferably there is at least 95%, more preferably at least 98%, homology to the sequence shown in Figure 1.

The above terms are synonymous with allelic variations of the sequences.

The term "complementary" means that the present invention also covers nucleotide sequences that can hybridise to the nucleotide sequence of the present invention.

The second strategy of the present invention will now be described in more detail by way of reference to further preferred aspects of the present invention.

According to a first preferred aspect of this aspect of the present invention there is provided an enzyme comprising the amino acid sequence shown in Figure 4, or a variant, homologue or fragment thereof.

According to a second preferred aspect of this aspect of the present invention there is provided a nucleotide sequence coding for the enzyme of the first aspect or a sequence that is complementary thereto.

According to a third preferred aspect of this aspect of the present invention there is provided a nucleotide sequence comprising the sequence shown in Figure 4, or a variant, homologue or fragment thereof or a sequence that is complementary thereto.

According to a fourth preferred aspect of this aspect of the present invention there is provided a construct comprising or expressing the invention according to any one of the earlier aspects.

According to a fifth preferred aspect of this aspect of the present invention there is provided a vector comprising or expressing the invention of any one of the earlier aspects.

According to a sixth preferred aspect of this aspect of the present invention there is provided a plasmid comprising or expressing the invention of any one of the earlier aspects.

According to a seventh preferred aspect of this aspect of the present invention there is provided a transgenic organism (or part thereof) comprising or expressing the invention according to any one of the earlier aspects.

Preferably in these preferred aspects of this aspect of the present invention the nucleotide sequence or the enzyme is that as defined in or is contained within or is expressed by the above-mentioned aspects of the present invention.

Other preferred aspects of this aspect of the present invention include methods of expressing or allowing expression or transforming any one of the nucleotide sequence, the construct, the plasmid, the vector, the cell, the tissue, the organ or the organism, as well as the products thereof.

Further preferred aspects of this aspect of the present invention include uses of the enzyme for preparing or treating foodstuffs, including animal feed.

A preferred aspect of this aspect of the present invention thus relates to the enzyme α-galactosidase and a nucleotide sequence coding for that enzyme.

In particular, the preferred aspect of the present invention relates to recombinant α-galactosidase.

In addition, the preferred aspect of the present invention relates to the use of that recombinant α-galactosidase to alter the mannose-to-galactose ratio of either an organism (or part thereof) and/or a mannose/galactose containing compound thereof, especially the mannose-to-galactose ratio of galactomannan.

One of the key advantages of the present invention is that by using the recombinant α-galactosidase it is possible to increase the mannose-to-galactose ratio of an organism (or part thereof) and/or a mannose/galactose containing compound thereof, in particular *in vivo* modified guar gum.

Other key advantages are that the recombinant enzyme can be prepared easily and in large quantities. Also, the nucleotide sequence can be used to change the *in vivo* ratio of mannose to galactose levels when inserted (preferably stably inserted) into the genome of an organism (or part thereof).

Preferably the nucleotide sequence is a DNA sequence. In a highly preferred embodiment the nucleotide sequence is a recombinant DNA sequence.

Preferably the nucleotide sequence is obtainable from deposit NCIMB 40831.

In a highly preferred embodiment the enzyme is expressed by use of recombinant DNA techniques.

Preferably the enzyme is expressed by a nucleotide sequence obtainable from deposit NCIMB 40831.

Preferably the organism is a plant.

More preferably the plant is a guar plant.

Preferably the mannose/galactose containing compound is guar gum.

The enzyme or nucleotide sequence(s) coding for same may be used *in vivo* in combination with one or more other enzymes or nucleotide sequence(s) coding for same, which enzymes or nucleotide sequence(s) coding for same are preferably prepared by use of recombinant DNA techniques. The α-galactosidase enzyme or nucleotide sequence(s) coding for same may also be used *in vitro*. The α-galactosidase enzyme or nucleotide sequence(s) coding for same may also be used with one or more other enzymes or nucleotide sequence(s) coding for same, which enzymes or nucleotide sequence(s) coding for same are preferably prepared by use of recombinant DNA techniques.

The terms "variant", "homologue" or "fragment" in relation to the enzyme include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) amino acid from or to the sequence providing the resultant amino acid sequence has α-galactosidase activity, preferably having at least the same activity of the enzyme shown in Figure 4. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant enzyme has α-galactosidase activity. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90%

homology to the sequence shown in Figure 4. More preferably there is at least 95%, more preferably at least 98%, homology to the sequence shown in the attached Figure 4.

The terms "variant", "homologue" or "fragment" in relation to the nucleotide sequence coding for the enzyme include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence providing the resultant nucleotide sequence codes for an enzyme having α-galactosidase activity, preferably having at least the same activity of the enzyme shown in Figure 4. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence codes for an enzyme having α-galactosidase activity. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to the sequence shown in Figure 4. More preferably there is at least 95%, more preferably at least 98%, homology to the sequence shown in Figure 4.

The above terms are synonymous with allelic variations of the sequences.

The term "complementary" means that the present invention also covers nucleotide sequences that can hybridise to the nucleotide sequence of the present invention.

The term "nucleotide" in relation to the present invention includes genomic DNA, cDNA, synthetic DNA, and RNA. Preferably it means DNA, more preferably cDNA for the coding sequence of the present invention.

The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes the nucleotide sequence directly or indirectly attached or fused to a promoter. An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the *Sh1*-intron or the ADH intron, intermediate the promoter and the nucleotide sequence. In each case, it is highly preferred that the terms do not cover the natural combination

of the wild type gene coding for the enzyme ordinarily associated with the wild type gene promoter and when they are both in their natural environment. One highly preferred embodiment of the present invention therefore relates to the nucleotide sequence of the present invention operatively linked to a heterologous promoter.

The construct may even contain or express a marker which allows for the selection of the genetic construct in, for example, a plant, such as guar, into which it has been transferred. Various markers exist which may be used, such as for example those encoding mannose-6-phosphate isomerase (especially for plants) or those markers that provide for herbicide or antibiotic resistance - e.g. resistance to G418, hygromycin, bleomycin, kanamycin and gentamycin.

The term "vector" includes expression vectors and transformation vectors.

The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression.

The term "transformation vector" means a construct capable of being transferred from one species to another - such as from an *E.coli* plasmid to an *Agrobacterium* to a plant.

The term "tissue" includes tissue *per se* and organ.

The term "organism" in relation to the present invention includes any organism that could comprise the nucleotide sequence coding for the enzyme according to the present invention and/or products obtained therefrom, and/or wherein the nucleotide sequence according to the present invention can be expressed when present in the organism.

Preferably the organism is a guar plant.

The term "transgenic organism" in relation to the present invention includes any organism that comprises the nucleotide sequence coding for the enzyme according to the present invention and/or products obtained therefrom, and/or wherein the nucleotide sequence according to the present invention can be expressed within the organism. Preferably the nucleotide sequence is incorporated in the genome of the organism.

Preferably the transgenic organism is a plant, more preferably a guar plant.

The transgenic organism of the present invention includes an organism comprising any one or more of the nucleotide sequences coding for the enzymes according to the present invention, constructs according to the present invention, vectors according to the present invention, plasmids according to the present invention, cells according to the present invention, tissues according to the present invention, or the products thereof, including combinations thereof. For example the transgenic organism can also comprise any one or more of the nucleotide sequences coding for the enzymes of the present invention under the control of one or more heterologous promoters.

In a highly preferred embodiment, the transgenic organism (or part thereof) does not comprise the combination of a promoter and the nucleotide sequence coding for the enzyme according to the present invention, wherein both the promoter and the nucleotide sequence are native to that organism (or part thereof) and are in their natural environment. Thus, in this highly preferred embodiment, the present invention does not cover the native nucleotide coding sequence according to the present invention in its natural environment when it is under the control of its native promoter which is also in its natural environment. In addition, in this highly preferred embodiment, the present invention does not cover the native enzyme according to the present invention when it is in its natural environment and when it has been expressed by its native nucleotide coding sequence which is also in its natural environment and when that nucleotide sequence is under the control of its native promoter which is also in its natural environment.

The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site in the Jacob-Mond theory of gene expression.

The promoter could additionally include one or more features to ensure or to increase expression in a suitable host. For example, the features can be conserved regions such as a Pribnow Box or a TATA box. The promoters may even contain other sequences to affect (such as to maintain, enhance, decrease) the levels of expression of the nucleotide sequence of the present invention. For example, suitable other sequences include the *Sh1*-intron or an ADH intron. Other sequences include inducible elements - such as temperature, chemical, light or stress inducible elements.

Also, suitable elements to enhance transcription or translation may be present. An example of the latter element is the TMV 5' signal sequence (see Sleat Gene 217 [1987] 217-225; and Dawson Plant Mol. Biol. 23 [1993] 97).

Thus, in one aspect, the nucleotide sequence according to the present invention is under the control of a promoter that allows expression of the nucleotide sequence. In this aspect, the promoter may be a cell or tissue specific promoter. If, for example, the organism is a plant then the promoter can be one that affects expression of the nucleotide sequence in any one or more of seed, stem, sprout, root and leaf tissues.

By way of example, the promoter for the nucleotide sequence of the present invention can be the α-Amy 1 promoter (otherwise known as the Amy 1 promoter, the Amy 637 promoter or the α-Amy 637 promoter) as described in PCT/EP95/02195.

Alternatively, the promoter for the nucleotide sequence of the present invention can be the α-Amy 3 promoter (otherwise known as the Amy 3 promoter, the Amy 351 promoter or the α-Amy 351 promoter) as described in PCT/EP95/02196. With the Amy 351 promoter it is possible to inactivate a part of it so that the partially inactivated promoter expresses the nucleotide sequence in a more specific manner such as in just one specific tissue type or organ. The term "inactivated" means partial

inactivation in the sense that the expression pattern of the promoter is modified but wherein the partially inactivated promoter still functions as a promoter. However, as mentioned above, the modified promoter is capable of expressing the nucleotide sequence in at least one (but not all) specific tissue of the original promoter. Examples of other partial inactivation of a promoter sequence (and not just necessarily that of the Amy 351 promoter) include altering the folding pattern of the promoter sequence, or binding species to parts of the nucleotide sequence, so that a part of the nucleotide sequence is not recognised by, for example, RNA polymerase. Another, and preferable, way of partially inactivating the Amy 351 promoter is to truncate it to form fragments thereof. Another way would be to mutate at least a part of the sequence so that the RNA polymerase can not bind to that part or another part.

Another modification is to mutate the binding sites for regulatory proteins for example the CreA protein known from filamentous fungi to exert carbon catabolite repression, and thus abolish the catabolite repression of the native promoter.

General teachings of recombinant DNA techniques may be found in Sambrook,J., Fritsch, E.F., Maniatis T. (Editors) Molecular Cloning. A laboratory manual. Second edition. Cold Spring Harbour Laboratory Press. New York 1989.

Even though the enzyme and the nucleotide sequence of the present invention are not disclosed in EP-B-0470145 and CA-A-2006454, those two documents do provide some useful background commentary on the types of techniques that may be employed to prepare transgenic plants according to the present invention. An adaption of some of these background teachings is now included in the following commentary.

The basic principle in the construction of genetically modified plants is to insert genetic information in the plant genome so as to obtain a stable maintenance of the inserted genetic material.

Several techniques exist for inserting the genetic information, the two main principles being direct introduction of the genetic information and introduction of the genetic information by use of a vector system. A review of the general techniques may be found in articles by Potrykus (Annu Rev Plant Physiol Plant Mol Biol [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27).

Thus, in one aspect, the present invention relates to a vector system which carries the nucleotide sequence or construct according to the present invention and which is capable of introducing the nucleotide sequence or construct into the genome of an organism, such as a plant.

The vector system may comprise one vector, but it can comprise two vectors. In the case of two vectors, the vector system is normally referred to as a binary vector system. Binary vector systems are described in further detail in Gynheung An et al. (1980), Binary Vectors, *Plant Molecular Biology Manual A3*, 1-19.

One extensively employed system for transformation of plant cells with a given promoter or nucleotide sequence or construct is based on the use of a Ti plasmid from *Agrobacterium tumefaciens* or a Ri plasmid from *Agrobacterium rhizogenes* An et al. (1986), *Plant Physiol.* 81, 301-305 and Butcher D.N. et al. (1980), *Tissue Culture Methods for Plant Pathologists,* eds.: D.S. Ingrams and J.P. Helgeson, 203-208.

Several different Ti and Ri plasmids have been constructed which are suitable for the construction of the plant or plant cell constructs described above. A non-limiting example of such a Ti plasmid is pGV3850.

The nucleotide sequence or construct of the present invention should preferably be inserted into the Ti-plasmid between the terminal sequences of the T-DNA or adjacent a T-DNA sequence so as to avoid disruption of the sequences immediately surrounding the T-DNA borders, as at least one of these regions appear to be essential for insertion of modified T-DNA into the plant genome.

As will be understood from the above explanation, if the organism is a plant, then the vector system of the present invention is preferably one which contains the sequences necessary to infect the plant (e.g. the *vir* region) and at least one border part of a T-DNA sequence, the border part being located on the same vector as the genetic construct.

Furthermore, the vector system is preferably an *Agrobacterium tumefaciens* Ti-plasmid or an *Agrobacterium rhizogenes* Ri-plasmid or a derivative thereof, as these plasmids are well-known and widely employed in the construction of transgenic plants, many vector systems exist which are based on these plasmids or derivatives thereof.

In the construction of a transgenic plant the nucleotide sequence or construct of the present invention may be first constructed in a microorganism in which the vector can replicate and which is easy to manipulate before insertion into the plant. An example of a useful microorganism is *E. coli*, but other microorganisms having the above properties may be used. When a vector of a vector system as defined above has been constructed in *E. coli*, it is transferred, if necessary, into a suitable *Agrobacterium* strain, e.g. *Agrobacterium tumefaciens*. The Ti-plasmid harbouring the nucleotide sequence or construct of the invention is thus preferably transferred into a suitable *Agrobacterium* strain, e.g. *A. tumefaciens*, so as to obtain an *Agrobacterium* cell harbouring the nucleotide sequence or construct of the invention, which DNA is subsequently transferred into the plant cell to be modified.

As reported in CA-A-2006454, a large amount of cloning vectors are available which contain a replication system in *E. coli* and a marker which allows a selection of the transformed cells. The vectors contain for example pBR 322, pUC series, M13 mp series, pACYC 184 etc.

In this way, the nucleotide or construct of the present invention can be introduced into a suitable restriction position in the vector. The contained plasmid is used for the transformation in *E.coli*. The *E.coli* cells are cultivated in a suitable nutrient medium

and then harvested and lysed. The plasmid is then recovered. As a method of analysis there is generally used sequence analysis, restriction analysis, electrophoresis and further biochemical-molecular biological methods. After each manipulation, the used DNA sequence can be restricted and connected with the next DNA sequence. Each sequence can be cloned in the same or different plasmid.

After each introduction method of the construct or nucleotide sequence according to the present invention in the plants the presence and/or insertion of further DNA sequences may be necessary. If, for example, for the transformation the Ti- or Ri-plasmid of the plant cells is used, at least the right boundary and often however the right and the left boundary of the Ti- and Ri-plasmid T-DNA, as flanking areas of the introduced genes, can be connected. The use of T-DNA for the transformation of plant cells has been intensively studied and is described in EP-A-120516; Hoekema, in: The Binary Plant Vector System Offset-drukkerij Kanters B.B., Alblasserdam, 1985, Chapter V; Fraley, et al., Crit. Rev. Plant Sci., 4:1-46; and An et al., EMBO J. (1985) 4:277-284.

Direct infection of plant tissues by *Agrobacterium* is a simple technique which has been widely employed and which is described in Butcher D.N. et al. (1980), *Tissue Culture Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208. For further teachings on this topic see Potrykus (Annu Rev Plant Physiol Plant Mol Biol [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27). With this technique, infection of a plant may be done on a certain part or tissue of the plant, i.e. on a part of a leaf, a root, a stem or another part of the plant.

Typically, with direct infection of plant tissues by *Agrobacterium* carrying the nucleotide sequence of the present invention, a plant to be infected is wounded, e.g. by cutting the plant with a razor or puncturing the plant with a needle or rubbing the plant with an abrasive. The wound is then inoculated with the *Agrobacterium*. The inoculated plant or plant part is then grown on a suitable culture medium and allowed to develop into mature plants.

When plant cells are constructed, these cells may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc.

Regeneration of the transformed cells into genetically modified plants may be accomplished using known methods for the regeneration of plants from cell or tissue cultures, for example by selecting transformed shoots using an antibiotic and by subculturing the shoots on a medium containing the appropriate nutrients, plant hormones, etc.

Further teachings on plant transformation may be found in EP-A-0449375.

Even further useful teachings on the transformation of guar plants can be found in Danish patent application No. 940662 (filed 10 June 1994).

Thus, the present invention relates to the use of a gene product (e.g. the PMI enzyme which is involved in the biosynthesis of mannose) to increase the mannose-to-galactose ratio of an organism (or part thereof) or mannose/galactose containing compound thereof. In addition, the present invention relates to that nucleotide sequence and the gene product it encodes.

The present invention is based on the surprising finding that it is possible to increase the mannose-to-galactose ratio of guar gum by the insertion of a gene or genes encoding a gene product or products which are involved in the biosynthesis of mannose/galactose containing compounds such as mannose-6-phosphate, namely PMI.

The findings of the present invention are in contrast to what would have been expected in view of the teachings of the art. In this regard, Edwards from analysis of unpurified preparations of developing endosperms from fenugreek and guar suggested that GDP-mannose may be a precursor for the biosynthesis of galactomannan (Edwards *et al* 1989, Reid and Edwards 1995). Through analysis of

various amounts of GDP-mannose in the reaction mixture the authors concluded that the control of the mannose-to-galactose ratio might reside at the level of specificity of the galactomannan-synthesizing glycosyltransferases themselves. This might suggest that the glycosyltransferases might be the critical target for genetic manipulation for the *in vivo* modification of guar gum.

Thus, in summation, this preferred aspect of the present invention relates to the insertion of a phosphomannose isomerase gene into a plant, preferably guar.

The rationale behind this strategy is based on our understanding that the mannose incorporated into the guar galactomannan is derived from GDP-mannose. However, the way in which the GDP-mannose is synthesized in guar is not known. The classical way is by isomerization of GDP-glucose but some preliminary data suggest that in some legumes at least the GDP-mannose might be synthesized by isomerization of fructose-6-phosphate to mannose-6-phosphate which is isomerized to mannose-1-phosphate which in turn is converted to GDP-mannose. But even if the latter pathway should be operative an increase in the PMI activity would not *per se* be expected to affect the galactomannan composition because firstly the fructose-6-phosphate to mannose-6-phosphate isomerization is a fully reversible reaction and secondly Reid and Edwards (1995) have predicted that the mannose-to-galactose ratio is determined by the specificity of the galactomannan synthesizing enzymes. With this strategy, and as will become apparent, we have obtained an increase of the mannose-to-galactose ratio - even when using a weakly expressed promoter in guar. But, even higher ratio levels could be achieved by using a stronger promoter.

The present invention is also based on the surprising finding that it is possible to increase the mannose-to-galactose ratio of guar gum by the insertion of a gene or genes encoding a gene product or products which are involved in the biosynthesis of guar or precursors therefor. As just mentioned, in one preferred aspect of the present invention the gene codes for PMI. In an alternative preferred aspect the gene codes for α-galactosidase, preferably coffee bean α-galactosidase or senna α-galactosidase, more preferably senna α-galactosidase.

The following sample comprising the PMI gene of the present invention was deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St. Machar Drive, Aberdeen, Scotland, United Kingdom, AB2 1RY on 9 November 1995:

E.coli K12 containing plasmid pPMI-60.

The deposit number is NCIMB 40774.

The present invention also covers nucleotide sequences obtainable from that deposit and the products encoded thereby.

The following sample comprising the α-galactosidase gene of the present invention was deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St. Machar Drive, Aberdeen, Scotland, United Kingdom, AB2 1RY on 28 November 1996:

E.coli K12 containing plasmid pT7-SEcDNA5.

The deposit number is NCIMB 40831.

The present invention also covers nucleotide sequences obtainable from that deposit and the products encoded thereby.

The present invention will now be described by way of example. In the following Examples reference is made to the accompanying figures in which:

Figure 1 shows the amino acid sequence of one enzyme according to the present invention and the sequence of one nucleotide sequence according to the present invention;

Figure 2 is a plasmid map of pcDNAII;

Figure 3 is a plasmid map of pSG-Man5;

Figure 4 shows the amino acid sequence of another enzyme according to the present invention and the sequence of another nucleotide sequence according to the present invention;

Figure 5 is a plasmid map of pPS48;

Figure 6 is a plasmid map of pPS48SEGAL;

Figure 7 is a plasmid map of pBKL4;

Figure 8 is a plasmid map of pBKL4SEGAL;

Figure 9 is a plasmid map of pPS48-GALIII; and

Figure 10 is a plasmid map of pBKL4GALIII.

IN VIVO MODIFICATION OF GUAR GUM USING PMI GENES

Cloning of a Phosphomannose Isomerase (PMI) Gene from Guar

A cDNA expression library represent mRNA from immature guar endosperm was constructed in the plasmid pcDNAII (Invitrogen Corporation) and transformed into the *E coli* strain Top10F~ (Invitrogen Corporation). The quality of the cDNA library was controlled by purification of plasmids from a number of separate colonies picked at random. Restriction enzyme digestion of the plasmids showed that all contained an insert of at least 500 bp.

The *E coli* strain CD1 man- contains an inactive PMI gene and is thus unable to metabolise mannose (Darzins *et al* 1985). This strain was used for the following complementation studies.

CD1 man- cells were made competent by the method of Hanahan (1985). A titer of 3-4 x 10^6 transformed cells/μg library plasmid cDNA was obtained. A similar titer was found when the cells were transformed with a Bluescript control plasmid.

Prior to the complementation studies, a number of control experiments were performed in which the transformed cells were plated onto selective medium containing M9-salts (Maniatis *et al* 1982) added 0.05 g/l leucine, 0.05 g/l methionine, 0.05 g/l threonine, 1.0 g/l thiamin-HC1, 50 mg/l ampicillin, 6.0 g/l mannose and 9.0 g/l agarose. This medium is hereafter called M9-SGP.

In one experiment CD1 man- competent cells were transformed with the *E coli* PMI gene under control of its native promoter (Mills and Guest 1984) and in another experiment the cells were transformed with the *E coli* PMI gene under the control of the plant promoter CaMV 35S (pSGMAN1, see Bojsen *et al* 1993). In both experiments, plating of the transformed cells on M9-SGP resulted in large numbers of large colonies. When competent CD1 man- cells were not transformed or transformed with a Bluescript control plasmid, no large colonies were obtained but a number of very small hardly visible colonies were observed. Thus the M9-SGP selective medium is suitable for selecting cells which contain an active PMI gene.

Competent CD1 man- cells were transformed with plasmid DNA isolated from the guar endosperm cDNA library. Transformed CD1 man- cells were plated onto the selective substrate M9-SGP. After incubation for two days at 37°C the majority of the plated cells appeared as very small colonies whereas less than 0.1% of the colonies were remarkably larger.

Plasmids from the larger colonies were purified and retransformed into competent CD1 man- cells. Twenty different retransformed colonies were assayed for PMI

activity in crude extracts. PMI activity was measured by the coupled enzymatic assay described by Gill *et al* (1986). One of the examined clones contained very high PMI activity. This clone is designated PMI-60 and is the subject of the deposit number NCIMB 40774.

In order to test whether the insert in PMI-60 originated from guar, genomic guar DNA was purified from leaves according to Dellaporta *et al* (1983), restriction enzyme digested, Southern blotted and probed with plasmid DNA derived from PMI-60 labelled with P-32 according to Feinberg and Vogelstein (1983). Hybridisation and washing were performed at 68°C, 6 x SSC and 68°C, 0.2 x SSC (Maniatis *et al* 1982), respectively. The PMI-60 probe hybridised to a number of fragments of the digested guar genomic DNA.

The insert in the plasmid of PMI-60 (called pPMI-60) was sequenced by dideoxy sequencing applying primer walking (Sanger and Coulson 1977), first by fluorescein labelled primers (reverse and universal) and subsequently by internal labelling using fluorescein-dATP.

The size of the cDNA clone was 1.66 kb of which the PMI gene covered 1.29 kb. The PMI start and stop codon was located at 0.09 kb and 1.38 kb, respectively. The insert contained a putative polyadenylation signal at 0.11 kb downstream from the stop codon and is poly-A terminated.

The identity of the insert was further characterised by comparison to other known PMI sequences. At 137-145 amino acids downstream from translation start a conserved region is found: DGNHKPEM which is considered to be involved in the active PMI site of phosphomannose isomerases (Coulin *et al* 1993).

Thus, the presence of the insert in PMI-60 results in rapid growth on selective mannose containing medium, high PMI activity and hybridisation to guar DNA. Furthermore, there is homology to some sequences from other PMI genes. These data demonstrate that the insert in PMI-60 is a PMI gene from guar.

This PMI sequence is the first PMI sequence ever cloned and sequenced from a plant.

Transformation using phosphomannose isomerase (PMI) gene

The following transformation studies show that it is possible to increase the mannose-to-galactose ratio of guar gum by the insertion of a phosphomannose isomerase (PMI) - such as that from *E.coli* or guar. The recombinant PMI catalyses the conversion of fructose-6-phosphate to mannose-6-phosphate or mannose-6-phosphate to fructose-6-phosphate depending on the substrate available.

Transformation of Guar

Transgenic guar plants were obtained by Agrobacterium tumefaciens mediated gene transfer as described by Jørsboe and Okkels (1994) - the contents of which are incorporated herein by reference. The preferred *Agrobacterium tumefaciens* strain was in these studies LBA 4404.

The insert in the T-DNA in the plasmid called pDO18 contained 3 genes (right border to left border): a β-glucuronidase (GUS) gene, a phosphomannose isomerase (PMI) gene and a neomycin phosphotransferase (NPTII) gene. The expression of each of the genes was driven by a 35S promoter described in detail on Bojsen *et al* (1993).

Analysis of Guar Gum by HPLC

Pure guar endosperms devoid of embryo and seed coat were prepared by hand and treated repeatedly with 70% ethanol during homogenization according to Edwards *et al* (1992). The ethanol precipitate was added 2 ml 2 N trifluor acetic acid (TFA) and hydrolysis was performed at 120°C for 2 hours. The TFA was removed by evaporation at 50°C. The dried precipitate was dissolved in 500 μl HPLC grade H_2O and 25 μl sample was applied onto an Aminex HPX-87P column (300 x 7.8 mm). The column was heated to 80°C in a column oven and eluted with H_2O. Elution of saccharides was followed with an RI-detector. Using this system mannose and

galactose was baseline separated. The respective peak areas were determined and the mannose-to-galactose ratios were calculated. Guar gum and LBG from Sigma were used as standards along with endosperm samples from non-transgenic guar plants. The mannose-to-galactose ratios obtained were 1.6:1 for both Sigma guar gum and non-transgenic endosperms, and 3.5:1 for Sigma LBG. These ratios are in excellent agreement with those generally accepted (Reid and Edwards 1995).

Mannose-to-Galactose Ratio in PMI-Transformed Guar

Several independent guar transformants harbouring the PMI gene were analysed with respect to mannose-to-galactose ratio in the endosperm galactomannan, see the table below.

Origin of Sample	Mannose-to-Galactose Ratio
Guar gum from Sigma	1.60
LBG from Sigma	3.44
Guar transformant 123-1	1.64
Guar transformant 123-2	1.63
Guar transformant 124-1	1.74

Each of the transformants increased the mannose-to-galactose ratio of guar gum.

Similar studies are done with the insertion of a phosphomannose isomerase (PMI) gene from guar. In this regard, it is believed that the guar PMI catalyses the conversion of fructose-6-phosphate to mannose-6-phosphate or mannose-6-phosphate to fructose-6-phosphate depending on the substrate available.

IN VIVO MODIFICATION OF GUAR GUM USING α-GALACTOSIDASE GENES

In the following examples it is shown that the insertion of an α-galactosidase gene into guar can result in the *in vivo* modification of guar gum.

Cloning and sequencing of senna α-galactosidase cDNA.

cDNA clones of α-galactosidase from senna endosperm were isolated by PCR as follows.

Total RNA was purified from senna endosperm according to the method of (Logemann *et al.* Anal. Biochem. 163 (1987) 16-20). Reverse transcription followed by PCR was done with the RT-PCR kit from Perkin Elmer and according to their protocol. Briefly, approximately 1 mg of total RNA and 1 mg oligo-dT was incubated at 42°C for 45 min with reverse transcriptase followed by incubation at 99°C for 5 min and 5°C for 5 min.

For the following PCR, two oligo nucleotides derived from the guar α-galactosidase cDNA sequence (Overbeeke et al. Plant Mol. Biol 13 (1989) 541-550) were used:

P1 (5'-CAACGGGGCTTGCTGCTTTAGG)

and

P2 (5'-GCCTATGTCA-GACCAGGATGC),

at positions 415-437 and 1248-1270, respectively in the guar α-galactosidase cDNA sequence.

The PCR conditions were: 1 min at 94°C, 2 min at 55°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C.

WEST

The PCR product was analysed by agarose gel electrophoresis, and a single product of 850 bp was obtained.

The DNA product, designated SEGAL, was cloned into the pT7Blue vector (Novagen) and the nucleotide sequence was determined using a Termo sequenase fluorescent cycle sequencing kit (Amersham) and an ALF DNA sequencer (Pharmacia).

The 3' and 5' ends of the senna α-galactosidase cDNA were obtained by the method called 3' and 5' RACE as previously described (Nielsen et al. Plant Mol. Biol 31 (1996) 539-552. Briefly, for 3' RACE, approximately 1mg of the above total RNA and 2.5 pmol of the primer:

Q_T (5' CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGC$(T)_{17}$)

was incubated at 42°C for 45 min with reverse transcriptase followed by incubation at 99°C for 5 min and 5°C for 5 min.

The cDNA was amplified by two rounds of PCR.

For the first PCR the downstream primer was:

Q_0 (5'-CCAGTGAGCAGAGTGACG),

and the upstream primer was:

3'GSP1 (5'-GTCCTCTGAGTGATAACAGAGTGG),

a gene specific primer derived from the 850 bp nucleotide sequence of the fragment SEGAL, position 1096-1118 in the senna α-galactosidase cDNA (Figure 4).

In the second PCR the downstream primer was:

Q_1 (5'-GAGGACTCGAGCTCAAGC)

and the upstream primer was:

3'GSP2 (5'-GGTGTTGTGGAATAGAAGTTCATC),

a gene specific primer derived from the 850 bp nucleotide sequence of the fragment SEGAL, position 1123-1146 in the senna α-galactosidase cDNA (Figure 4).

The PCR conditions for the first PCR were: 1 min at 94°C, 2 min at 60°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C. For the second PCR the conditions were: 1 min at 94°C, 2 min at 50°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C.

The PCR product after the second PCR was analysed by agarose gel electrophoresis, and a single product of 530 bp was obtained.

The DNA product, designated 3'SEGAL, was cloned into the pT7Blue vector (Novagen) and the nucleotide sequence was determined using a Termo sequenase fluorescent cycle sequencing kit (Amersham) and an ALF DNA sequencer (Pharmacia).

For 5' RACE the 5'RACE system from Gibco BRL was employed with 3 gene specific primers constructed from the 850 bp nucleotide sequence of the fragment SEGAL. Briefly, approximately 1 mg of the same total RNA as used above and 2.5 pmol of the gene specific primer:

5'GSP1 (5'-TTGCACCTTGGTCTTCATGTCC),

position 561-582 in the senna α-galactosidase cDNA (Figure 4) was incubated for 70°C for 10 min followed by addition of reverse transcriptase and incubating at 42°C for 30 min, 70°C for 15 min and the addition of RNaseH and incubating further 10

min at 55°C.

The cDNA was dC-tailed according to the protocol of Gibco BRL. The tailed cDNA was subjected to two rounds of PCR.

In the first PCR an upstream ANKER primer (5'-GGCCACGCGTC-GACTAGTACGGGGGGGGGG) was used with the gene specific downstream primer:

5'GSP2 (5-CATAGCTTTACTGCATGTTTGGTTTCC),

position 510-536 in the senna α-galactosidase cDNA (Figure 4).

In the second PCR an upstream UNI primer:

(5'-GGCCACGCGTCGACTAGTACG)

was used with the gene specific downstream primer:

5'GSP3 (5'-CAGCCAGAGCCTTAATTCCTGAAGG),

position 438-462 in the senna α-galactosidase cDNA (Figure 4).

The PCR conditions for the first PCR were: 1 min at 94°C, 2 min at 51°C, 2 min at 72°C for 10 cycles followed by 1 min at 94°C, 2 min at 59°C, 2 min at 72°C for 25 cycles followed by 10 min at 72°C.

For the second PCR the conditions were: 1 min at 94°C, 2 min at 59°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C.

The PCR product after the second PCR was analysed by agarose gel electrophoresis, and a single product of 480 bp was obtained.

The DNA product, designated 5'SEGAL, was cloned into the pT7Blue vector (Novagen) and the nucleotide sequence was determined using a Termo sequenase fluorescent cycle sequencing kit (Amersham) and an ALF DNA sequencer (Pharmacia).

The complete nucleotide sequence of the senna α-galactosidase cDNA, combined from the 3 PCR clones, 5'SEGAL, SEGAL and 3'SEGAL, obtained above, consisting of 1630 bp is shown in Figure 4. Analysis of the nucleotide sequence reveals (Figure 4) an open reading frame encoding 406 amino acid residues with the first methionine at nucleotide position 93 and a translation stop codon at nucleotide position 1311. The deduced amino acid sequence is shown above the nucleotide sequence in Fig 4.

Construction of a plant transformation vector containing the senna α-galactosidase.

An expression vector comprising the coding sequence for senna α-galactosidase was constructed as follows.

The vector pPS48 (Figure 5) was constructed by inserting the 0.75 kb cauliflower mosaic virus (CaMV) 35S RNA promoter (E35S) containing a duplication of the -90 to -420 region (Kay et al. Science 236 (1987) 1299-1302), the 0.21 kb fragment containing the CaMV 35S RNA polyadenylation sequence (Odell et al. Nature 313 (1985) 810-812) and a synthetic oligonucleotide linker (*Pst*I-*BamH*I-*Sma*I-*Sac*I-*Sal*I-*Sph*I) into pUC8 (Vieira and Messing, Gene 19 (1982) 259-268).

The three DNA fragments, 5'SEGAL, SEGAL and 3'SEGAL, were linked together by PCR to reconstitute a clone of senna α-galactosidase cDNA, containing the coding sequence and most of the non translated 5' end.

The fragment 5' SEGAL was re-amplified using the primers 5'GSP3 and B255,

(5'- ATTGGATCCACTCACAC-GTATACACTACAC)

containing a *BamHI* site plus the nucleotides position 14-36 in the senna α-galactosidase cDNA (Figure 4).

The fragment 3'SEGAL was re-amplified using the primers 3'GSP2 and B254,

(5'-TTAGCATGCCCTTGGGATTGTATT-TCCTC),

containing a *SphI* site and nucleotides position 1353-1373 in the senna α-galactosidase cDNA (Figure 4).

These PCR fragment were mixed with the fragment SEGAL, and the linked sequence was amplified by using the flanking primers B254 and B255.

The PCR conditions were: For 5'SEGAL, 1 min at 94°C, 2 min at 57°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C. For 3'SEGAL, 1 min at 94°C, 2 min at 48°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C. For the last PCR to link the three fragments, 1 min at 94°C, 2 min at 58°C, 2 min at 72°C for 35 cycles followed by 10 min at 72°C.

The PCR product of the last PCR was analysed by agarose gel electrophoresis, and a single product of 1380 bp was obtained.

The DNA product was cloned into the pT7Blue vector (Novagen) and the nucleotide sequence was determined using a Termo sequenase fluorescent cycle sequencing kit (Amersham) and an ALF DNA sequencer (Pharmacia).

The DNA fragment was isolated by digestion with BamHI and SphI and cloned into the expression vector pPS48, resulting in pPS48SEGAL (Figure 6).

A plant transformation vector comprising the expression cassette for senna α-galactosidase was constructed as follows.

The vector pBKL4 (Figure 7) was constructed from pBI121 (Clontec Laboratories) by deletion of the NPTII cassette and by insertion of a synthetic oligonucleotide linker (*EcoR*I-*Cla*I-*Sal*I-*Hind*III-*Spe*I-*Kpn*-*BamH*I) and a new NPTII cassette containing the NPTII gene flanked by a CaMV 35S RNA promoter (Odell et al. Nature 313 (1985) 810-812) and a polyadenylation sequence from the octopine synthase gene (Caplan et al. Science 222 (1983) 815-821) between the GUS cassette and the left border.

The senna α-galactosidase expression cassette was excised from pPS48SEGAL and inserted into pBKL4.

The resulting plasmid pBKL4SEGAL (Figure 8) was transformed into *Agrobacterium tumefaciens* strain LBA4404 (Hockema et al. Nature 303 (1983) 179-180) for plant transformation.

Transformation of guar with senna α-galactosidase

The α-galactosidase gene from senna was transformed into guar using *Agrobacterium tumefaciens* transformation, as described in detail by Jørsboe and Okkels (1994) - the contents of which are incorporated herein by reference.

Analysis of the mannose-to-galactose ratio of endosperms from guar plants transformed with the senna α-galactosidase gene.

Pure endosperms of guar plants transformed with the senna α-galactosidase gene were analysed by HPLC after hydrolysis in 2 N trifluor acetic acid (see above for details).

The results, which are shown in the following Table, are derived from the analysis of endosperms from 4 independent guar transformants harbouring the senna α-galactosidase gene.

The results are given as % increases of the mannose-to-galactose ratios relative to untransformed control endosperms.

Transformant	Relative Increase of the Mannose-to-Galactose Ratio
26-5-12-1	6.5%
29-5-08-1	5.3%
39-5-45-1	18.9%
40-5-02-1	10.6%

Construction of a plant transformation vector containing the coffee bean α-galactosidase.

An expression vector comprising the coding sequence for coffee bean α-galactosidase was constructed as follows.

The vector pPS48 (Figure 5) was constructed by inserting the 0.75 kb cauliflower mosaic virus (CaMV) 35S RNA promoter (E35S) containing a duplication of the -90 to -420 region (Kay et al. Science 236 (1987) 1299-1302), the 0.21 kb fragment containing the CaMV 35S RNA polyadenylation sequence (Odell et al. Nature 313 (1985) 810-812) and a synthetic oligonucleotide linker (*Pst*I-*Bam*HI-*Sma*I-*Sac*I-*Sal*I-*Sph*I) into pUC8 (Vieira and Messing, Gene 19 (1982) 259-268).

A DNA fragment containing the coffee bean α-galactosidase coding sequence was isolated from the plasmid pCR-BZ (Zhu and Goldstein Gene 140 (1994) 227-231) by polymerase chain reaction (PCR) employing an upstream primer:

5'-TTGGATCCACCCAAAA-GCTGGTGCTCC

(position 15-35 in the coffee bean α-galactosidase cDNA sequence) and a downstream

primer:

5'-TTAGCATGCCTGTTAATCACTGTGGG

(position 1229-12446 in the coffee bean α-galactosidase cDNA sequence) resulting in a 1,2 kb DNA fragment containing a *BamHI* site in the 5' end of the α-galactosidase gene and a *SphI* site in the 3' end.

The DNA fragment was cloned into the pT7Blue vector (Novagen) and the nucleotide sequence was determined using a Termo sequenase fluorescent cycle sequencing kit (Amersham) and an ALF DNA sequencer (Pharmacia).

The DNA fragment was isolated by digestion with *BamHI* and *SphI* and cloned into the vector pPS48, resulting in pPS48-GALIII (Figure 9).

A plant transformation vector comprising the expression cassette for coffee bean α-galactosidase was constructed as follows.

The vector pBKL4 (Figure 7) was constructed from pBI121 (Clontec Laboratories) by deletion of the NPTII cassette and by insertion of a synthetic oligonucleotide linker (*EcoRI-ClaI-SalI-HindIII-SpeI-Kpn-BamHI*) and a new NPTII cassette containing the NPTII gene flanked by a CaMV 35S RNA promoter (Odell et al. Nature 313 (1985) 810-812) and a polyadenylation sequence from the octopine synthase gene (Caplan et al. Science 222 (1983) 815-821) between the GUS cassette and the left border.

The coffee bean α-galactosidase expression cassette was excised from pPS48-GALIII by digestion with *XbaI* and inserted into *SpeI* digested *pBKL4*.

The resulting plasmid pBKL4-GALIII (Figure 10) was transformed into *Agrobacterium tumefaciens* strain LBA4404 (Hockema et al. Nature 303 (1983) 179-180) for plant transformation.

Transformation of guar with coffee bean α-galactosidase

The α-galactosidase gene from coffee bean was transformed into guar using *Agrobacterium tumefaciens* transformation, as described in detail by Jørsboe and Okkels (1994).

Analysis of the mannose-to-galactose ratio of endosperms from guar plants transformed with the coffee bean α-galactosidase gene.

Pure endosperms of guar plants transformed with the coffee bean α-galactosidase gene were analysed by HPLC after hydrolysis in 2 N trifluor acetic acid (see above for details). The results, which are shown in the following Table, are derived from the analysis of endosperms from 4 independent guar transformants harbouring the coffee bean α-galactosidase gene. The results are given as % increases of the mannose-to-galactose ratios relative to un-transformed control endosperms.

Transformant	Relative Increase of the Mannose-to-Galactose Ratio
4-2-4-2	4.2%
4-2-5-5	4.0%
4-2-8-2	6.0%
17-3-6-1	4.1%

Discussion of *in vivo* modification of guar gum

In the two examples presented above, it is shown that the insertion of an α-galactosidase gene into guar can result in the *in vivo* modification of guar gum as evidenced by the fact that the mannose-to-galactose ratios of the transgenic galactomannans were higher than the non-transgenic control galactomannans.

In addition to the example with an phosphomannose isomerase gene (see above), these are the first data ever presented on the *in vivo* modification of a plant cell wall storage polysaccharide.

SUMMATION

The results show that it is possible to form transgenic plants that are capable of forming guar gum such that the mannose-to-galactose ratio is affected.

In this regard, the PMI transformed lines, the coffee bean α-galactosidase transformed lines, and the senna α-galactosidase transformed lines yielded a higher ratio of mannose to galactose than the non-transformed lines. For example, the coffee bean α-galactosidase transformants had a ratio of mannose to galactose of up to 1.75 - compared to 1.65 for the non-transformed lines - and that some of the senna α-galactosidase transformants had a ratio of mannose to galactose of even up to 2. These results are highly surprising.

As will be apparent to those skilled in the art, the extent to which the *in vivo* modification occurs depend on the activity of the galactomannan related enzymes encoded by the genes transformed into guar. Thus, substitution or modification of the promoters or other regulatory nucleotide or amino acid sequences may lead to *in vivo* modifications of galactomannans different from those described in the above examples.

Other modifications of the present invention will be apparent to those skilled in the art.

As mentioned above, Figure 1 presents both a nucleotide sequence coding for a PMI enzyme and the amino acid sequence of that PMI enzyme. For the avoidance of doubt, that nucleotide sequence can be referred to as SEQ ID No. 1 and that amino acid sequence can be referred to as SEQ ID No. 2.

As mentioned above, Figure 4 presents both a nucleotide sequence coding for an α-galactosidase enzyme and the amino acid sequence of that α-galactosidase enzyme. For the avoidance of doubt, that nucleotide sequence can be referred to as SEQ ID No. 3 and that amino acid sequence can be referred to as SEQ ID No. 4.

REFERENCES

Bojsen *et al* (1993) PCT patent application WO 94/20627

Bulpin P V, Gidley M J, Jeffcoat R, Underwood D R. Development of a biotechnological process for the modification of galactomannan polymers with plant α-galactosidase. Carbohydr. Polymers 1990, 12:155.

Coulin F, Magnenat E, Proutfood A E I, Payton M A and Wells T N C (1993). Biochemistry 32 14139-14144

Darzins *et al* 1985. J Bacteriol 161 249-257

Dellaporta S L, Wood J and Hicks J B (1983). A plant DNA minipreparation: Version II. Plant Molecular Biology Reporter 1 No 4, 19-21.

Edwards M, Bulpin P V, Dea I C M, Reid J S G. Biosynthesis of legume-seed galactomannan *in vitro*. Planta. 1989, 178:41.

Feinberg A P and Vogelstein B (1983). A technique for radio-labelling DNA restriction endonuclease fragments to high specific activity. Analytical Biochemistry 132, 6-13.

Gill *et al* (1986) J Bacteriol 167, 611-615

Hanahan D (1985). Techniques for Transformation of *E coli*. In: Glover D M (ed) DNA cloning vol I - a practical approach. IRL Press, Oxford. pp 109-135. ISBN: 0-947946-18-7.

Jørsboe M and Okkles F T (1994) Transformation of guar. PCT patent application PCT/DK95/00221.

McCleary B V, Critchley P, Bulpin B V. Processing of polysaccharides. European patent application, 1983, 0 121 960.

Maniatis T, Fritsch E F, Sambrook J (1982). Molecular cloning a laboratory manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York. ISBN: 0-87969-136-0.

Mills and Guest. 1984. Gene 32 41-48.

Overbeeke N, Fellinger A J, Hughes S G. Production of guar alpha-galactosidase by hosts transformed by recombinant DNA methods. European patent application, 1986, 0 255 153.

Reid J S G, Edwards M E. Galactomannans and other cell wall storage polysaccharides in seeds. In: food polysaccharides and their application. Ed: A M Stephen, Marcel Dekker, 1995, pp 155-186.

Sanger F, Micklen S and Coulson A R (1977). Proc. Natl. Acad. Sci. USA 74 5463-5467.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 28, lines 1 — 9

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☐

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution *(including postal code and country)*

23 St. Machar Drive
Aberdeen
Scotland
AB2 1RY
United Kingdom

Date of deposit: 9 NOVEMBER 1995

Accession Number: NCIMB 40774

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)*

This information is continued on an additional sheet ☐

In respect of those designations in which a European patent is sought, and any other designated state having equivalent legislation, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer Mrs. H. Fransz	Authorized officer

Form PCT/RO/134 (July 1992)

WEST

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 28, lines 14 - 21

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution (*including postal code and country*)

23 St. Machar Drive
Aberdeen
Scotland
AB2 1RY
United Kingdom

Date of deposit	Accession Number
28 NOVEMBER 1996	NCIMB 40831

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) — This information is continued on an additional sheet ☐

In respect of those designations in which a European patent is sought, and any other designated state having equivalent legislation, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer Mrs. H. Fransz	Authorized officer

Form PCT/RO/134 (July 1992)

WEST

PO Box 17,
DK 1001 Copennagen K,
Denmark.

RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT
issued pursuant to Rule 7.1 by the
INTERNATIONAL DEPOSITARY AUTHORITY
identified at the bottom of this page

I. IDENTIFICATION OF THE MICROORGANISM

Identification reference given by the DEPOSITOR:	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY:
Escherichia coli K12 CD-PMI-60 (CD1 man containing plasmid pPMI-60)	NCIMB 40774

II. SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION

The microorganism identified under I above was accompanied by:

[ ] a scientific description

[X] a proposed taxonomic designation

(Mark with a cross where applicable)

III. RECEIPT AND ACCEPTANCE

This International Depositary Authority accepts the microorganism identified under I above, which was received by it on 9 November 1995 (date of the original deposit)[1]

IV. RECEIPT OF REQUEST FOR CONVERSION

The microorganism identified under I above was received by this International Depositary Authority on (date of the original deposit) and a request to convert the original deposit to a deposit under the Budapest Treaty was received by it on (date of receipt of request for conversion)

V. INTERNATIONAL DEPOSITARY AUTHORITY

Name: NCIMB Ltd 23 St Machar Drive Aberdeen S Address: UK AB2	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Date: 13 November 1995

1 Where Rule 6.4(d) applies, such date is the date on which the status of international depositary authority was acquired.

Form BP/4 (sole page)

WEST

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

Danisco Biotechnology,
Langebrogade 1,
PO Box 17,
DK 1001 Copenhagen K,
Denmark.

NAME AND ADDRESS OF THE PARTY TO WHOM THE VIABILITY STATEMENT IS ISSUED

INTERNATIONAL FORM

VIABILITY STATEMENT
issued pursuant to Rule 10.2 by the INTERNATIONAL DEPOSITARY AUTHORITY identified on the following page

I. DEPOSITOR	II. IDENTIFICATION OF THE MICROORGANISM
Name: AS ABOVE Address:	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: NCIMB 40774 Date of the deposit or of the transfer:[1] 9 November 1995

III. VIABILITY STATEMENT

The viability of the microorganism identified under II above was tested on 10 November 1995 [2]. On that date, the said microorganism was

[X][3] viable

[ ][3] no longer viable

[1] Indicate the date of the original deposit or, where a new deposit or a transfer has been made, the most recent relevant date (date of the new deposit or date of the transfer).

[2] In the cases referred to in Rule 10.2(a)(ii) and (iii), refer to the most recent viability test.

[3] Mark with a cross the applicable box.

Form BP/9 (first page)

WEST

IV. CONDITIONS UNDER WHICH THE VIABILITY TEST HAS BEEN PERFORMED[4]	
V. INTERNATIONAL DEPOSITARY AUTHORITY	
Name: NCIMB Ltd Address: 23 St Machar Drive Aberdeen Scotland UK AB2 1RY	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Date: 13 November 1995

[4] Fill in if the information has been requested and if the results of the test were negative.

Form BP/9 (second and last page)

WEST

CLAIMS

1. An *in vivo* modification process that affects the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof.

2. An *in vivo* modification process that affects the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, the *in vivo* modification process comprising expressing a nucleotide sequence coding for a gene product, which gene product has an effect on:

   (a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

   (b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the nucleotide sequence is not a natural nucleotide sequence to the organism (or part thereof).

3. An *in vivo* modification process that affects the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, the *in vivo* modification process comprising allowing a gene product that is capable of having an effect on:

   (a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

   (b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

to have an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the gene product has not been expressed by a nucleotide sequence that is a natural nucleotide sequence to the organism (or part thereof).

4. Use of a nucleotide sequence to affect *in vivo* the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or of a mannose/galactose containing compound thereof, wherein the nucleotide sequence encodes a gene product which has an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the nucleotide sequence is not a natural nucleotide sequence to the organism (or part thereof).

5. Use of a gene product to affect *in vivo* the mannose-to-galactose ratio of either an organism (or part thereof) capable of producing a mannose/galactose containing compound or a mannose/galactose containing compound thereof, wherein the gene product has an effect on:

(a) the mannose-to-galactose ratio of mannose and galactose components of a mannose/galactose containing compound; and/or

(b) the mannose-to-galactose ratio of mannose and galactose precursors for a mannose/galactose containing compound;

and wherein the gene product is not expressed by a nucleotide sequence that is a natural nucleotide sequence to the organism (or part thereof).

6. The invention according to any one of claims 1 to 5 wherein the mannose/galactose containing compound is galactomannan.

7. The invention according to any one of claims 1 to 6 wherein the organism capable of producing a mannose/galactose containing compound is a guar plant.

8. The invention according to any one of claims 1 to 6 wherein the *in vivo* mannose-to-galactose ratio of the organism (or part thereof) or mannose/galactose containing compound thereof is higher than that of the guar plant or the galactomannan thereof.

9. The invention according to any one of claims 1 to 8 wherein the *in vivo* mannose-to-galactose ratio of the organism (or part thereof) or mannose/galactose containing compound thereof is substantially similar to that of the locust bean or the galactomannan thereof.

10. The invention according to any one of claims 2 to 9 wherein the gene product is at least one gene product useful in the biosynthesis of GDP-mannose.

11. The invention according to claim 10 wherein the gene product is the protein shown in Figure 1 or is a variant, homologue or derivative thereof.

12. The invention according to claim 10 or claim 11 wherein the gene product is coded by the nucleotide sequence shown in Figure 1 or is a variant, homologue or derivative thereof and/or is obtainable from NCIMB 40774.

13. The invention according to any one of claims 2 to 9 wherein the gene product is an α-galactosidase enzyme.

14. The invention according to claim 13 wherein the gene product is the protein shown in Figure 4 or is a variant, homologue or derivative thereof.

15. The invention according to claim 13 or claim 14 wherein the gene product is coded by the nucleotide sequence shown in Figure 4 or is a variant, homologue or derivative thereof and/or is obtainable from NCIMB 40831.

16. An enzyme comprising the amino acid sequence shown in Figure 1, or a variant, homologue or fragment thereof.

17. A nucleotide sequence coding for the enzyme of claim 16 or a sequence that is complementary thereto and/or is obtainable from NCIMB 40774.

18. A nucleotide sequence according to claim 17 wherein the nucleotide sequence is a DNA sequence.

19. A nucleotide sequence comprising the sequence shown in Figure 1, or a variant, homologue or fragment thereof or a sequence that is complementary thereto and/or is obtainable from NCIMB 40774.

20. A construct comprising or expressing the invention according to any one of claims 16 to 19.

21. A vector comprising or expressing the invention of any one of claims 16 to 20.

22. A plasmid comprising or expressing the invention of any one of claims 16 to 21.

23. A transgenic organism (or part thereof) comprising or expressing the invention according to any one of claims 16 to 22.

24. A transgenic organism (or part thereof) according to claim 23 wherein the organism is a guar plant.

25. The invention according to claim 10 wherein the gene product is expressed by or is the invention according to any one of claims 16 to 24.

26. An enzyme comprising the amino acid sequence shown in Figure 4, or a variant, homologue or fragment thereof.

27. A nucleotide sequence coding for the enzyme of claim 26 or a sequence that is complementary thereto and/or is obtainable from NCIMB 40831.

28. A nucleotide sequence according to claim 27 wherein the nucleotide sequence is a DNA sequence.

29. A nucleotide sequence comprising the sequence shown in Figure 4, or a variant, homologue or fragment thereof or a sequence that is complementary thereto and/or is obtainable from NCIMB 40831.

30. A construct comprising or expressing the invention according to any one of claims 26 to 29.

31. A vector comprising or expressing the invention of any one of claims 26 to 30.

32. A plasmid comprising or expressing the invention of any one of claims 26 to 31.

33. A transgenic organism (or part thereof) comprising or expressing the invention according to any one of claims 26 to 32.

34. A transgenic organism (or part thereof) according to claim 33 wherein the organism is a guar plant.

35. The invention according to claim 9 wherein the gene product is expressed by or is the invention according to any one of claims 26 to 34.

36. A mannose/galactose containing compound when prepared by the process of any one of claims 1 to 3 or any claim dependent thereon.

37. A foodstuff comprising a mannose/galactose containing compound according to claim 36.

38. A composition comprising a mannose/galactose containing compound according to claim 36 admixed with another polysaccharide.

39. A composition comprising a mannose/galactose containing compound according to claim 36 admixed with any one or more of xanthan, carrageenan and agarose.

40. A method for preparing a composition or a foodstuff comprising mixing the mannose/galactose containing compound according to claim 36 with another suitable ingredient.

41. A process substantially as described herein.

42. A nucleotide sequence substantially as described herein and with reference to Figure 1.

43. An enzyme substantially as described herein and with reference to Figure 1.

44. A nucleotide sequence substantially as described herein and with reference to Figure 4.

45. An enzyme substantially as described herein and with reference to Figure 4.

Done on DNA sequence PMI60.

DE SGP60.R
DE MODIFIED BY INVERT COMPLEMENTATION FROM 60R.
DE MODIFIED BY INVERT COMPLEMENTATION FROM R06.
DE MODIFIED BY INVERT COMPLEMENTATION FROM 60RNY.
DE MODIFIED BY INVERT COMPLEMENTATION FROM YNR06.

Total number of bases is: 1556.
Analysis done on the complete sequence.
Done on (absolute) phase(s): 1.
Using the Universal genetic code.

```
         10        20        30        40        50        60
          |         |         |         |         |         |
CCAGTGTGGTGGAATTCCCTTCCCAATTCCCATTTCCTTTCCGTCACTAACCTTCCTTTT

         70        80        90       100       110       120
          |         |         |         |         |         |
GCATCAATTCACTCGCTCGCTATCACCATAGCCATGGAGGAAAAGGCTCATGCCACCAAG
                                   M  E  E  K  A  H  A  T  K

        130       140       150       160       170       180
          |         |         |         |         |         |
CAGCACCAGCATCTTCAGAGGCTTCATTGCTCCGTCAAGAACTACGATTGGGGTTTACCT
 Q  H  Q  H  L  Q  R  L  H  C  S  V  K  N  Y  D  W  G  L  P

        190       200       210       220       230       240
          |         |         |         |         |         |
GGTCGAATTTCCAATGTCGCGAGGCTCTACGCGCTCAATTCCGGGTCCCAATTCCATCCG
 G  R  I  S  N  V  A  R  L  Y  A  L  N  S  G  S  Q  F  H  P

        250       260       270       280       290       300
          |         |         |         |         |         |
GATGAGCCTTACGCGGAACTTTGGATGGGTACCCACGATTCAGAGCCATCGTTTCTTGTC
 D  E  P  Y  A  E  L  W  M  G  T  H  D  S  E  P  S  F  L  V

        310       320       330       340       350       360
          |         |         |         |         |         |
TCCAATGGTGCTCAAAGGGTCACCCTCAAGGCTTGGATTTCCCAGAACCCTGATGTGCTA
```

Fig 1

S N G A Q R V T L K A W I S Q N P D V L

370 380 390 400 410 420

GGTGAAAAGGTTCTTCAGAAATGGGGTTGTGATCTCCCTTTCTTGTTCAAGGTGCTGTCT

G E K V L Q K W G C D L P F L F K V L S

430 440 450 460 470 480

GTGGGGAAAGCCTTATCCATACAGGCTCACCCGGATAAGGAGTTGGCTAGGACTCTGCAT

V G K A L S I Q A H P D K E L A R T L H

490 500 510 520 530 540

AAATTGCAGCCTAATCTCTATAAGGATGGCAATCACAAACCTGAGATGGCTCTCGCAATG

K L Q P N L Y K D G N H K P E M A L A M

550 560 570 580 590 600

ACAGATTTCGAAGCTCTTTGTGGATTCATCACTCCTAAGGAGCTCAAGGCTGTGCTTCAT

T D F E A L C G F I T P K E L K A V L H

610 620 630 640 650 660

ACTGTTCCTGAAGTTGTTGAACTGGTTGGCGCTGCAAATGCAAACCTAATCTTACAAACT

T V P E V V E L V G A A N A N L I L Q T

670 680 690 700 710 720

AGAGATCAGGATGGTGAAGAAAAGGTAAAACCTGTTTTGCAGACAGTATTCACCCACCTC

R D Q D G E E K V K P V L Q T V F T H L

730 740 750 760 770 780

ATGTCAGCTAGTAAAGAGAAAGTAACTGATGCAGTAAACAGATTGAAAAGTCGTCTGCAT

M S A S K E K V T D A V N R L K S R L H

790 800 810 820 830 840

AAGGAAAGTGAGGTGAGGCAGTTGACAGATAAGGAGCAGCTAGTGCTGCGATTGGAAAAG

K E S E V R Q L T D K E Q L V L R L E K

850 860 870 880 890 900

WEST

```
CAATACCCAGCTGATGTTGGTGTCATAGCTGCCTTCTTTCTTAACCTTGTAAAACTCAGT
 Q  Y  P  A  D  V  G  V  I  A  A  F  F  L  N  L  V  K  L  S

       910       920       930       940       950       960
CCTGGTGAAGCATTGTTCCTAGGGGCAAACGAACCACACGCATATCTACATGGGGAGTGC
 P  G  E  A  L  F  L  G  A  N  E  P  H  A  Y  L  H  G  E  C

       970       980       990      1000      1010      1020
ATTGAATGCATGGCAACTTCAGACAATGTTGTGCGAGCTGGCCTATCTCCCAAACACAGA
 I  E  C  M  A  T  S  D  N  V  V  R  A  G  L  S  P  K  H  R

      1030      1040      1050      1060      1070      1080
GATGTCCAGACCCTTTGTTCCATGCTTACATACAAACAGGGTTCTCCAGAGATATTGCCG
 D  V  Q  T  L  C  S  M  L  T  Y  K  Q  G  S  P  E  I  L  P

      1090      1100      1110      1120      1130      1140
GGAGTTCCTCTAAATCCATATGTAAAAAAATACACCCCGCCATTCAAGGAATTTGAGATT
 G  V  P  L  N  P  Y  V  K  K  Y  T  P  P  F  K  E  F  E  I

      1150      1160      1170      1180      1190      1200
GATAGATGTATTCTTCCTCAAGGGAAAACAGTGGTGTTCCCAGCAGTGCCAGGTCCTTCT
 D  R  C  I  L  P  Q  G  K  T  V  V  F  P  A  V  P  G  P  S

      1210      1220      1230      1240      1250      1260
ATCTTTTTGGTCACAGTTGGGGAAGGAGTGATGAATACAGGATCACCATCCAAAGAACAA
 I  F  L  V  T  V  G  E  G  V  M  N  T  G  S  P  S  K  E  Q

      1270      1280      1290      1300      1310      1320
GTTGTCTCTGAAGGTGATGTCCTTTTTGCAGCTGCTTACACTGAGATTAGTGTTACTAGT
 V  V  S  E  G  D  V  L  F  A  A  A  Y  T  E  I  S  V  T  S

      1330      1340      1350      1360      1370      1380
GCATCTGAGTTGCATCTGTACAGAACGGGAGTTAATAGCAGATTTTTTCAGGCTTCCTAA
 A  S  E  L  H  L  Y  R  T  G  V  N  S  R  F  F  Q  A  S  -
```

WEST

```
     1390      1400      1410      1420      1430      1440
        |         |         |         |         |         |
GTTTCAACTGAACGTTCCTTTTTTAGGGGCCATGGCTTTTGAGAGTCAAACTGTGCAATT

     1450      1460      1470      1480      1490      1500
        |         |         |         |         |         |
GTGTGTTATATATGTAATAGTATTAGTAGCAGTTTAGTGTACATGGAGTAATAACTTAGT

     1510      1520      1530      1540      1550
        |         |         |         |         |
CTCAGAGAAAAGTTCATCCAAAGAGAGGTCGGTTGCCAAAAAAAAAAAAAAAAAAA
```

WEST

Comments for pcDNA II
DNA Length: 3013 nucleotides
First nucleotide: +1
Date Created: 4/12/1990 (2h:57m:34s)
Last modified: 3/23/1992 (10h:7m:5s)

This is the sequence of pcDNA II
T7 promoter bases 2993-3012
Polylinker bases 10-122
Sp6 promoter 136-152
Ampicillin resistance bases 1331-2191
F1 origin bases 2377-2832
Lac Z gene bases 2932-300

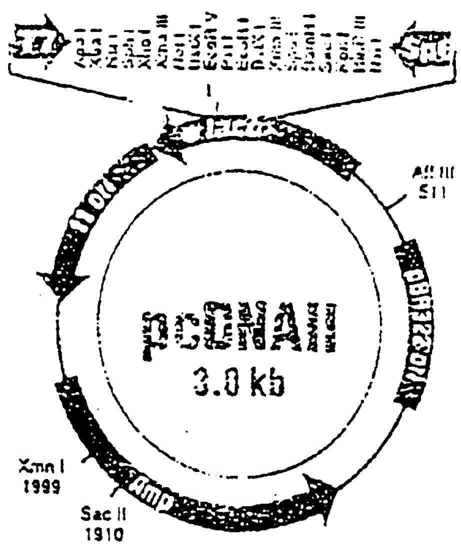

Fig 2

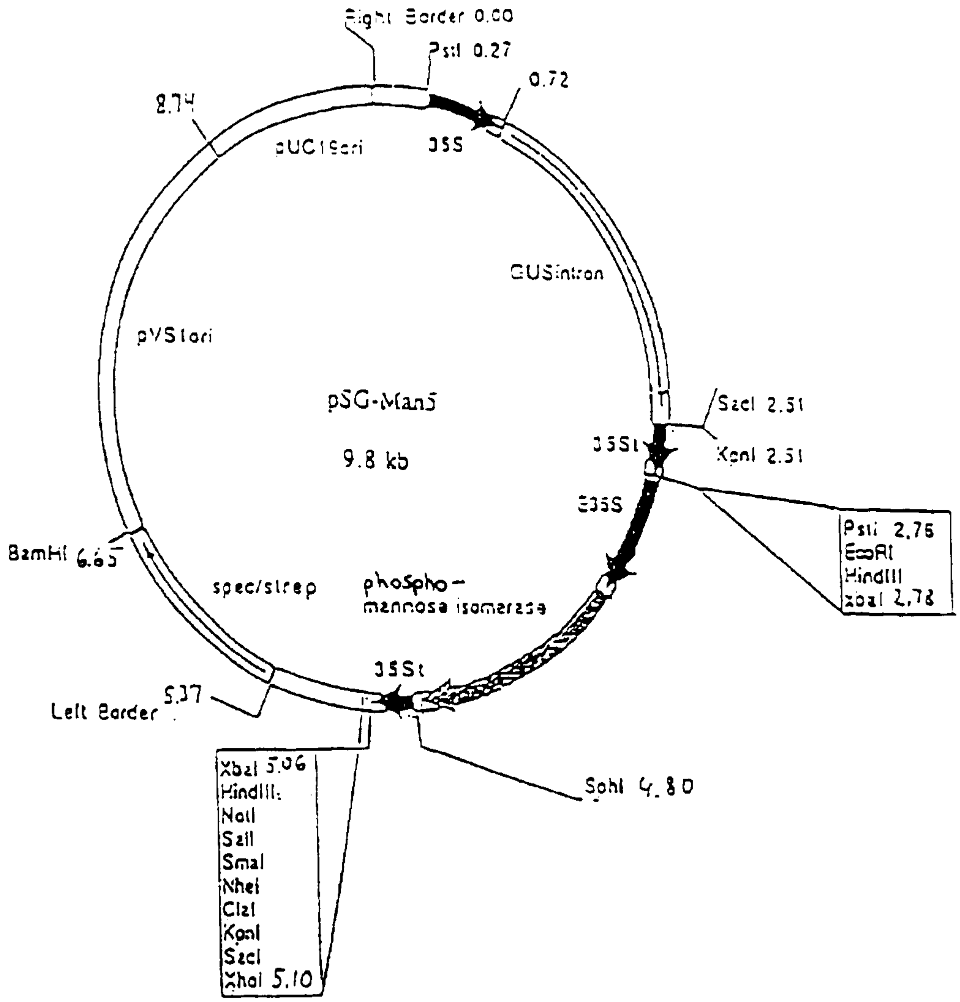
Right Border 0.00
PstI 0.27
0.72
8.74
pUC19ori
35S
GUSintron
pVS1ori
pSG-Man5
9.8 kb
SacI 2.51
KpnI 2.51
35St
35S
PstI 2.76
EcoRI
HindIII
XbaI 2.78
BamHI 6.65
spec/strep
phospho-
mannose isomerase
35St
Left Border 5.37
XbaI 5.06
HindIII
NotI
SalI
SmaI
NheI
ClaI
KpnI
SacI
XhoI 5.10
SphI 4.80

Fig 3

WEST

```
*******************************************
* TRANSLATION OF A NUCLEIC ACID SEQUENCE *
*******************************************

Done on DNA sequence SECDNA5.

Total number of bases is: 1630.

        10        20        30        40        50        60
         |         |         |         |         |         |
GGGGGGGGGGATAACACTCACACGTATACACTACACACACACTCATATATATATAGAGAG

        70        80        90       100       110       120
         |         |         |         |         |         |
AGAGAGAGAGAGAGAGATTTATATAGAAAGAAATGGAGAAAATGATGATGTGGGCAAAGG
                                  M  E  K  M  M  M  W  A  K

       130       140       150       160       170       180
         |         |         |         |         |         |
TTGTGTTGTGCTTGTTTTGGGTCTTGAATGCTTCTAATTGTTCAGGTCGCTTGTTGAACA
V  V  L  C  L  F  W  V  L  N  A  S  N  C  S  G  R  L  L  N

       190       200       210       220       230       240
         |         |         |         |         |         |
CAATTGGCAATGATCACAACAACATCCATGGAAGACTACTTCTTGGAAATGGACTTGGAA
T  I  G  N  D  H  N  N  I  H  G  R  L  L  L  G  N  G  L  G

       250       260       270       280       290       300
         |         |         |         |         |         |
ACACTCCTCCCATGGGATGGAATAGCTGGAACCACTTTCAGTGTGACATTAATGAGGAGA
N  T  P  P  M  G  W  N  S  W  N  H  F  Q  C  D  I  N  E  E

       310       320       330       340       350       360
         |         |         |         |         |         |
TGGTTCGAGAAACAGCTGATGCAATGGTGTCAACGGGTCTTGCATCTTTGGGGTACGAAT
M  V  R  E  T  A  D  A  M  V  S  T  G  L  A  S  L  G  Y  E

       370       380       390       400       410       420
         |         |         |         |         |         |
ACGTCAATTTAGATGATTGCTGGGCTGAACTTAACCGAGACTCTAAGGGAAATATGGTTC
Y  V  N  L  D  D  C  W  A  E  L  N  R  D  S  K  G  N  M  V

       430       440       450       460       470       480
         |         |         |         |         |         |
CTAGTGCTTCAAAATTTCCTTCAGGAATTAAGGCTCTGGCTGATTATGTTCATAGCAAAG
P  S  A  S  K  F  P  S  G  I  K  A  L  A  D  Y  V  H  S  K

       490       500       510       520       530       540
         |         |         |         |         |         |
GATTGAAGTTTGGGGTTTATTCTGATGCTGGAAACCAAACATGCAGTAAAGCTATGCCTG
G  L  K  F  G  V  Y  S  D  A  G  N  Q  T  C  S  K  A  M  P

       550       560       570       580       590       600
         |         |         |         |         |         |
GATCACTTGGACATGAAGACCAAGGTGCAAAAACATTTGCTTCCTGGGGGGTTGATTTCT
G  S  L  G  H  E  D  Q  G  A  K  T  F  A  S  W  G  V  D  F
```

Fig. 4

```
        610       620       630       640       650       660
          |         |         |         |         |         |
TGAAGTATGACAATTGTAACAACAATGATATAAGCCCAAGAAATAGGTATCCAAAAATGT
L  K  Y  D  N  C  N  N  N  D  I  S  P  R  N  R  Y  P  K  M

        670       680       690       700       710       720
          |         |         |         |         |         |
CTGAAGCTCTAGCAAATTCTGGAAGGGCAATCTTCTTCTCTATGTGTGAATGGGGATCAG
S  E  A  L  A  N  S  G  R  A  I  F  F  S  M  C  E  W  G  S

        730       740       750       760       770       780
          |         |         |         |         |         |
AAGACCCTGCACTATGGGCCAAAAGTGTGGGAAATAGTTGGAGAACAACTGGAGATATTG
E  D  P  A  L  W  A  K  S  V  G  N  S  W  R  T  T  G  D  I

        790       800       810       820       830       840
          |         |         |         |         |         |
AAGATAAGTGGGAAAGTATGGCATCTATTGCTGACCAAAATGACAAATGGGCATCTTATG
E  D  K  W  E  S  M  A  S  I  A  D  Q  N  D  K  W  A  S  Y

        850       860       870       880       890       900
          |         |         |         |         |         |
CTGGGCCTGGAGGGTGGAATGATCCTGATATGCTTGAAGTTGGAAATGGAGGCATGACAA
A  G  P  G  G  W  N  D  P  D  M  L  E  V  G  N  G  G  M  T

        910       920       930       940       950       960
          |         |         |         |         |         |
CAGAAGAATATCGTTCTCATTTTAGCATATGGGCATTAGCTAAGGCTCCTTTGTTGATTG
T  E  E  Y  R  S  H  F  S  I  W  A  L  A  K  A  P  L  L  I

        970       980       990      1000      1010      1020
          |         |         |         |         |         |
GTTGTGATGTTCGATCAATGGATGGCGCAACATACGGACTGCTAAGCAACAAGGAAGTTA
G  C  D  V  R  S  M  D  G  A  T  Y  G  L  L  S  N  K  E  V

       1030      1040      1050      1060      1070      1080
          |         |         |         |         |         |
TTGCAGTAAACCAAGACAGTCTAGGAGTCCAAGGAAAGAAGGTGAAAAGTGATGCTGGTT
I  A  V  N  Q  D  S  L  G  V  Q  G  K  K  V  K  S  D  A  G

       1090      1100      1110      1120      1130      1140
          |         |         |         |         |         |
TGGAGGTTTGGGCAGGTCCTCTGAGTGATAACAGAGTGGCAGTGGTGTTGTGGAATAGAA
L  E  V  W  A  G  P  L  S  D  N  R  V  A  V  V  L  W  N  R

       1150      1160      1170      1180      1190      1200
          |         |         |         |         |         |
GTTCATCAAAAGCTACTGTGACTGCATCCTGGTCTGACATAGGGCTTGAAAAAGGAAAGG
S  S  S  K  A  T  V  T  A  S  W  S  D  I  G  L  E  K  G  K

       1210      1220      1230      1240      1250      1260
          |         |         |         |         |         |
TGGTCACTGCAAAAGATTTATGGGAGCACACTACAAAAGCATCAGTTTCAGGACAAATTT
V  V  T  A  K  D  L  W  E  H  T  T  K  A  S  V  S  G  Q  I
```

```
        1270      1280      1290      1300      1310      1320
           |         |         |         |         |         |
CTGCAGATATAGATTCACATGCTTGTAAGATGTATGTTCTGACTCCCAATTAAGGCAGAC
S  A  D  I  D  S  H  A  C  K  M  Y  V  L  T  P  N  -

        1330      1340      1350      1360      1370      1380
           |         |         |         |         |         |
AGGAAGGTGATGAAAGCCAAGGTTTTAAGGAAGAGGAAATACAATCCCAAGGGTTCCAAA

        1390      1400      1410      1420      1430      1440
           |         |         |         |         |         |
AGAGGGTGGGAGAAAATAAAACATGGGAGTTTATTTTTCAATAAGGAATATATATGGAA

        1450      1460      1470      1480      1490      1500
           |         |         |         |         |         |
ATAATTTAAATAAAGATAAGAAATTCGCCCTGAGATTTCTATACATTATTTGTATTATTG

        1510      1520      1530      1540      1550      1560
           |         |         |         |         |         |
GAGCATTGAGATTTGAGACCTTTTGATTTCAATTCAATAATTATACCGCAAATGTTTCTC

        1570      1580      1590      1600      1610      1620
           |         |         |         |         |         |
TATATTTTCCATAATTATGGTGATTATAATAAATGTTAAATTTGCTTGATGAAAAAAAAA

        1630
           |
AAAAAAAAAA
```

WEST

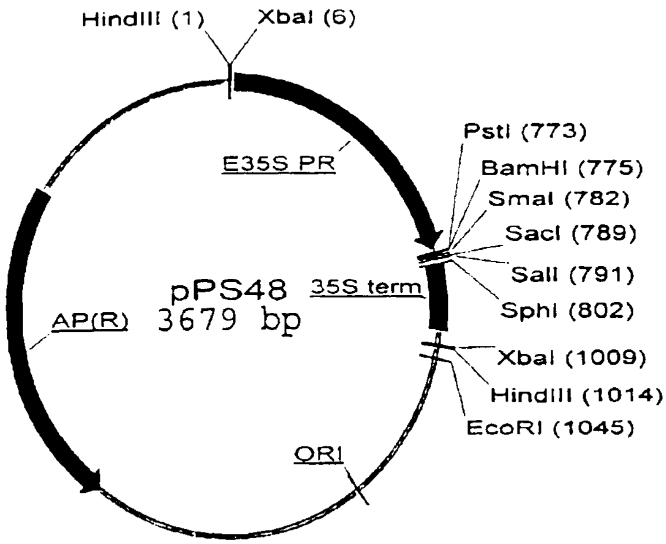

Fig. 5

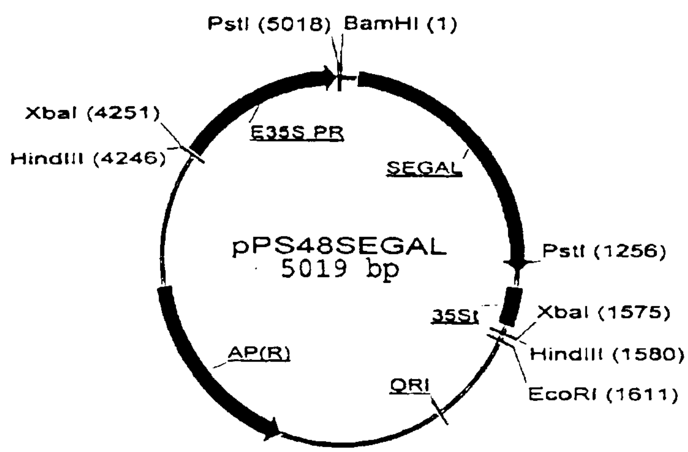

Fig 6.

WEST

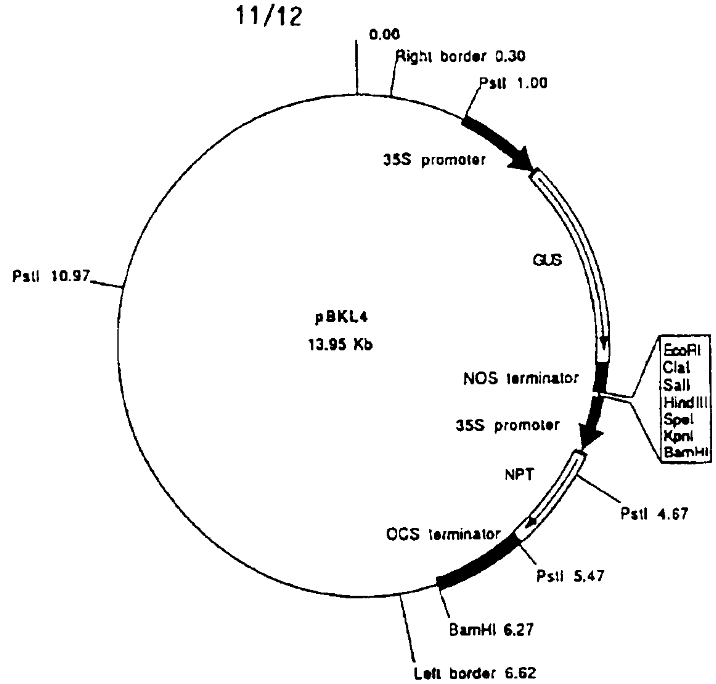
0.00
Right border 0.30
PstI 1.00
35S promoter
GUS
PstI 10.97
pBKL4
13.95 Kb
NOS terminator
EcoRI
ClaI
SalI
HindIII
SpeI
KpnI
BamHI
35S promoter
NPT
PstI 4.67
OCS terminator
PstI 5.47
BamHI 6.27
Left border 6.62

Fig. 7.

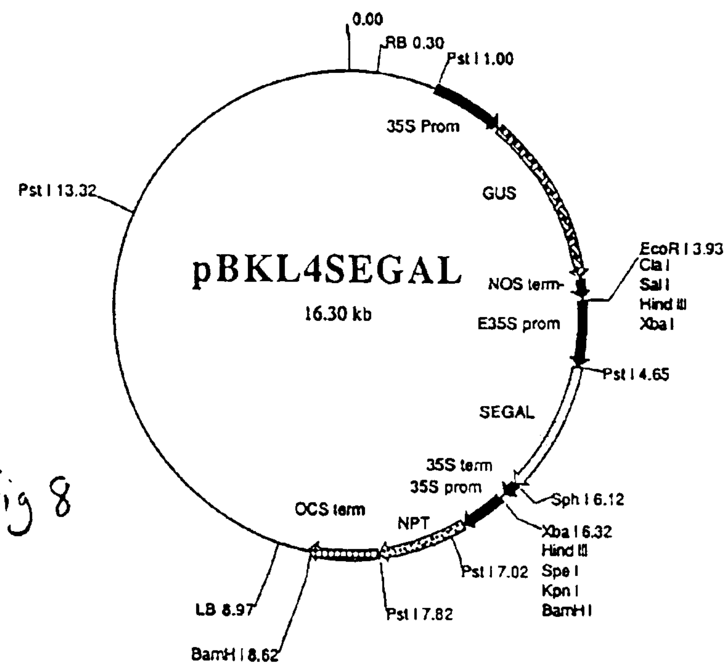
0.00
RB 0.30
Pst I 1.00
35S Prom
GUS
Pst I 13.32
pBKL4SEGAL
16.30 kb
NOS term
EcoR I 3.93
Cla I
Sal I
Hind III
Xba I
E35S prom
Pst I 4.65
SEGAL
35S term
35S prom
Sph I 6.12
Xba I 6.32
Hind III
Spe I
Kpn I
BamH I
OCS term
NPT
Pst I 7.02
Pst I 7.82
LB 8.97
BamH I 8.62

Fig 8

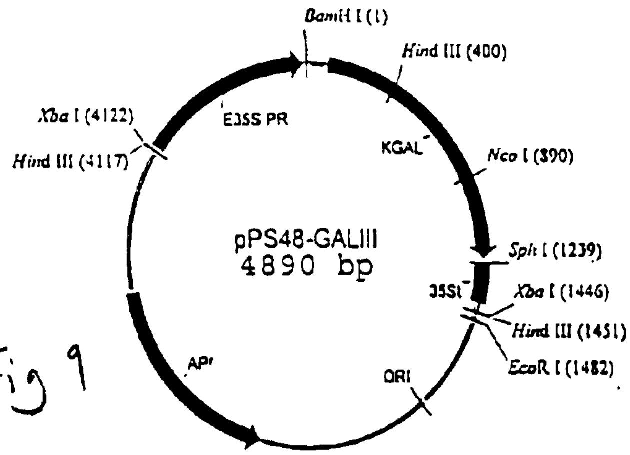

Fig 9

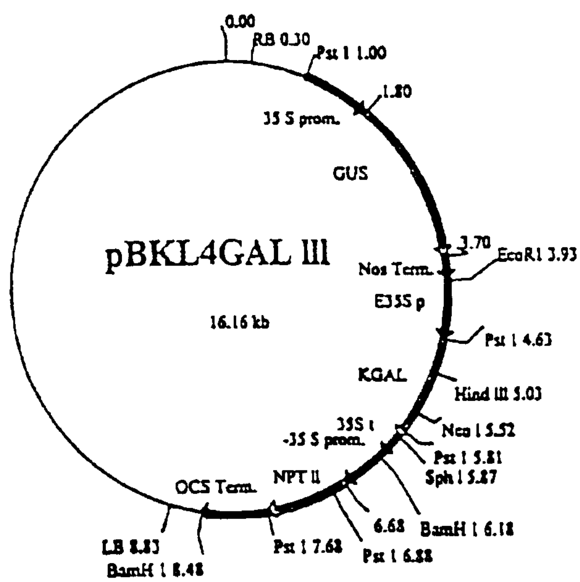

Fig 10

WEST